PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 9/10, 15/54, 15/62
C12N 15/80, 15/85

A1

(11) International Publication Number: **WO 92/12238**

(43) International Publication Date: 23 July 1992 (23.07.92)

(21) International Application Number: PCT/US91/09784

(22) International Filing Date: 30 December 1991 (30.12.91)

(30) Priority data:
635,756 4 January 1991 (04.01.91) US

(71) Applicant: THE BOARD OF REGENTS, UNIVERSITY OF TEXAS SYSTEM [US/US]; 201 West 7th Street, Austin, TX 78701 (US).

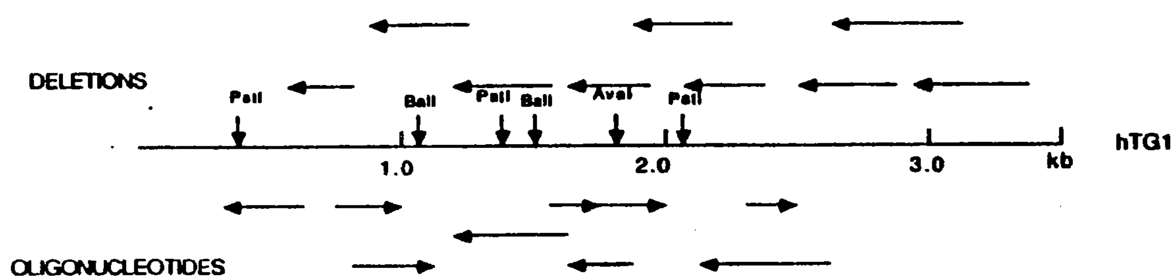
(72) Inventors: DAVIES, Peter, J., A. ; 5503 Cheltenham Drive, Houston, TX 77096 (US). STEIN, Joseph, P. ; 1971 Chard Road, Cazenovia, NY 13035 (US).

(74) Agents: PARMELEE, Steven, W. et al.; Townsend and Townsend, One Market Plaza - 2000 Steuart Tower, San Francisco, CA 94105 (US).

(81) Designated States: AT, AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH, CH (European patent), CI (OAPI patent), CM (OAPI patent), CS, DE, DE (European patent), DK, DK (European patent), ES, ES (European patent), FI, FR (European patent), GA (OAPI patent), GB, GB (European patent), GN (OAPI patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU, LU (European patent), MC (European patent), MG, ML (OAPI patent), MN, MR (OAPI patent), MW, NL, NL (European patent), NO, PL, RO, RU ,SD, SE, SE (European patent), SN (OAPI patent), TD (OAPI patent), TG (OAPI patent).

Published
With international search report.

(54) Title: CLONING AND EXPRESSION OF TISSUE TRANSGLUTAMINASES



(57) Abstract

Human and murine tissue transglutaminases are cloned, sequenced and expressed. The tissue transglutaminases herein are useful for, *inter alia*, therapeutic wound repair, stabilizing food preparations, and markers for identifying agents which act as agonists or antagonists of cellular apoptosis.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	RU	Russian Federation
CG	Congo	KP	Democratic People's Republic of Korea	SD	Sudan
CH	Switzerland	KR	Republic of Korea	SE	Sweden
CI	Côte d'Ivoire	LI	Liechtenstein	SN	Senegal
CM	Cameroon	LK	Sri Lanka	SU	Soviet Union
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
DE	Germany	MC	Monaco	TG	Togo
DK	Denmark			US	United States of America

CLONING AND EXPRESSION OF TISSUE TRANSGLUTAMINASES

Background of the Invention

Transglutaminases are a group of calcium dependent enzymes that catalyze the crosslinking of proteins by promoting the formation of ϵ-(γ-glutaminyl)lysine isopeptide bonds between protein-bound glutamine and lysine residues. These enzymes are believed to be widely distributed in nature, as the crosslinks are found in both prokaryotic and eukaryotic cells. Although different transglutaminases appear to be very similar in substrate specificity, several distinct forms of the enzymes have been identified. See generally, Folk, Ann. Rev. Biochem. 49:517-531 (1980).

Transglutaminase-mediated protein crosslinking reactions have been implicated in both normal and pathological processes in mammalian cells and tissues. The crosslink may act to maintain some forms of protein structure, such as in the terminal differentiation of epidermal cell layers and in other cellular architecture. An intracellular transglutaminase known as epidermal or Type I transglutaminase has been isolated and cloned from rabbit epithelial cells (Floyd and Jetten, Mol. Cell. Biol. 9:4846-4851 (1989)), and a transglutaminase has been isolated and cloned from guinea pig liver cells (Ikura et al., Biochem. 27: 2898-2905 (1988)). Other transglutaminases include hair follicle transglutaminase, keratinocyte transglutaminase, and prostate transglutaminase (Wilson et al., Fed. Proc. 38:1809 (1979)). Lee et al., Prep. Biochem. 16:321-335 (1986) have described the purification of a transglutaminase from human erythrocytes. These transglutaminases have been shown to be distinct from a plasma transglutaminase, Factor XIII, an enzyme whose primary function appears to be stabilizing fibrin clots. Factor XIII has also been purified, cloned, and sequenced. (Ichinose, et al.,

BNSDOCID: <WO 9212238A1 I >

Biochem. 25:6900-6906 (1986), Takahashi, et al., Proc. Natl. Acad. Sci. U.S.A. 83:8018-8023 (1986)).

The transglutaminases have been employed for crosslinking purposes in a variety of fields. Certain microbial transglutaminases have found use in food technology to add texture to processed foods, particularly fish and cheese. Others have been used in enzyme catalyzed fluorescent labeling of proteins, in the introduction of cleavable crosslinks, and in the solid phase reversible removal of specific proteins from biological systems. Factor XIII preparations have been proposed for a variety of therapeutic uses, such as the treatment of subarachnoid hemorrhage and inflammatory bowel disease.

Presently, a plasma derived Factor XIII is available as a fibrin sealant, but, as with most plasma-derived products, carries an inherent risk of viral contamination. Further, Factor XIII and certain other transglutaminases are zymogens, requiring some form of activation to become catalytically active. And, as each transglutaminase has a restricted range of substrates, their activity may be limited in certain applications. Accordingly, what is needed in the art are methods for producing by recombinant means human and murine transglutaminases, particularly those transglutaminases which do not require activation to become catalytically active. The present invention fulfills these and other related needs.

Summary of the Invention

The present invention provides the ability to produce human and murine tissue transglutaminases and polypeptides or fragments thereof by recombinant means, preferably in cultured eukaryotic cells. The expressed transglutaminase may or may not have the biological activity of the native enzyme, depending on the intended use. Accordingly, isolated and purified polynucleotides are described which code for the transglutaminases and fragments thereof, where the polynucleotides may be in the form of DNA, such as cDNA or

genomic DNA, or RNA. Based on these sequences probes may be designed for hybridization to identify these and related genes or transcription products thereof which encode human and murine tissue transglutaminases.

In related embodiments the invention concerns DNA constructs which comprise a transcriptional promoter, a DNA sequence which encodes the transglutaminase or fragment thereof, and a transcriptional terminator, each operably linked for expression of the enzyme or enzyme fragment. The constructs are preferably used to transform or transfect host cells, preferably eukaryotic cells, more preferably yeast or mammalian cells. For large scale production the expressed transglutaminase may be isolated from the cells by, for example, immunoaffinity purification.

Nucleic acid sequences which encode the transglutaminases of the invention and the recombinant transglutaminases themselves can also be used to develop compounds which can alter transglutaminase-associated apoptosis of a eukaryotic cell. Compounds may be screened for agonistic or antagonistic effects on transglutaminase-mediated metabolism in the host cell.

Brief Description of the Figures

Fig. 1 illustrates restriction maps of the mouse macrophage tissue transglutaminase cDNA clones, where the black portion of the boxes represents coding sequence for tissue transglutaminase, the white portion represents 3'-untranslated sequences, and the restriction sites are represented as A=Alu I, B=BamH I, and N=Nco I;

Fig. 2 illustrates the sequencing strategy for the human cDNA insert in clone hTG1;

Fig. 3 illustrates nucleotide sequences of human endothelial (SEQ. ID. No. 1) and mouse macrophage (SEQ. ID. No. 3) tissue transglutaminases and their predicted amino acid sequences (SEQ. ID. NO. 2 and SEQ. ID. NO. 4, respectively), where the wavy lines indicate the amino acid sequence of the

BNSDOCID: <WO 9212238A1 I >

pentapeptide containing the active site cysteine residue, the nucleotide sequence corresponding to a putative polyadenylation signal in the mouse sequence is located at the position 3452-3457, and the nucleotide sequence derived from mouse heart cDNA library has been underlined;

Fig. 4 illustrates the identification of human tissue transglutaminase mRNA by blot hybridization, analyzing 10 μg of mRNA from HUVEC; and

Fig. 5 illustrates the identification of mouse transglutaminase mRNA by blot hybridization, analyzing 10 μg of mRNA from each of the following tissues: L, liver; S, spleen; K, kidney; T, testis; H, heart; Lu, lung; Tm, thymus; and B, brain.

Description of the Preferred Embodiments

Tissue transglutaminase (or transglutaminase II) is an enzyme that catalyzes the crosslinking of protein-bound glutamine and primary amines, such as lysine residues. The present invention provides isolated nucleotide sequences of human tissue transglutaminase, thereby providing for the ultimate expression of human tissue transglutaminase polypeptides. Recombinant DNA expression systems provide convenient means for obtaining large quantities of human tissue transglutaminases in relatively pure form. The invention also provides cloned nucleotide sequences of murine tissue transglutaminase.

The invention also provides recombinant human and murine tissue transglutaminase polypeptides and fragments thereof having transglutaminase activity. By polypeptides and fragments is meant to include sequences of amino acids up to entire proteins, which have at least about 85% homology, preferably at least 90%, and more preferably at least about 95% or more homology to the amino acid sequences of the murine or human sequences of the invention, as shown in Fig. 3 and SEQ. ID. Nos. 1-4. As will be appreciated by those skilled in the art, the invention also includes those polypeptides having

slight variations in amino acid sequences or other properties. Such variations may arise naturally as allelic variations (e.g., due to genetic polymorphism) or may be produced by human intervention (e.g., by mutagenesis of cloned DNA sequences), such as induced point, deletion and insertion mutants.

Nucleic acid sequences encoding human tissue transglutaminase as described herein can be cloned from a variety of human cell sources that express the enzyme. Preferred sources include human umbilical vein endothelial cells and retinoic acid stimulated macrophages. Useful nucleic acid sequences in this regard include mRNA, genomic DNA and cDNA. For expression, cDNAs are generally preferred because they lack introns that may interfere with expression. To obtain a human tissue transglutaminase clone, a human endothelial cell cDNA library is screened with, e.g., labeled probes from random primed mouse macrophage transglutaminase sequences, which probes preferably span the enzyme's active site and/or putative calcium binding site. To obtain the mouse tissue transglutaminase clone, an oligo-dT primed cDNA library can be constructed with polyA$^+$ RNA purified from mouse peritoneal macrophages stimulated with retinoic acid. The library is screened with, e.g., polyclonal antibodies to guinea pig liver tissue transglutaminase and/or labeled RNA probes. Partial clones may be used as probes in additional screening until the complete coding sequence is obtained. If necessary, partial clones are joined in the correct reading frame to construct the complete coding sequence. Joining is achieved by digesting clones with appropriate restriction endonucleases and joining the fragments enzymatically in the proper orientation. Depending on the fragments and the particular restriction endonucleases chosen, it may be necessary to remove unwanted DNA sequences through a "loop out" process of deletion mutagenesis or through a combination of restriction endonuclease cleavage and mutagenesis. It is preferred that the resultant sequence be in the form of a continuous open reading frame, that is, that it lack intervening sequences (introns). The sequence of one exemplary mouse clone described herein, TGHZ3, includes 29 nucleotides of 5'-untranslated

sequence and 1,775 nucleotides of coding sequence and is shown in Fig. 3 (SEQ. ID. NO. 3).

A human cDNA transglutaminase clone isolated as described herein includes the entire 5'-untranslated sequence, as determined by primer extension analysis, the coding domain, and 1,058 nucleotides of 3'-untranslated sequence, as shown in Fig. 3 (SEQ. ID. NO. 1). This clone lacks a consensus polyadenylation sequence and is slightly shorter than the 3.6 kb full length transcript, as determined by Northern blot analysis of human endothelial cell RNA, suggesting that it lacks approximately 300 bp of 3'-untranslated sequence. The identity of the human tissue transglutaminase clone is confirmed by, for example, in vitro translation. As described further below, clone hTG-1 encodes a polypeptide that migrates at Mr 80,000 on SDS-polyacrylamide gels. Its deduced molecular weight is 77,253. The active site Cys residue was determined to be at position 277 as shown in Fig. 3.

With the nucleotide and deduced amino acid sequences of human tissue transglutaminase provided herein, genomic or cDNA sequences encoding tissue transglutaminase may be obtained from libraries prepared from other cells and tissues according to known procedures. For instance, using oligonucleotide probes derived from human endothelial transglutaminase sequences, generally of at least about fourteen nucleotides and up to twenty-five or more nucleotides in length, DNA sequences encoding transglutaminase of other tissues and/or mammalian species may be obtained. If partial clones are obtained, it is necessary to join them in proper reading frame to produce a full length clone, using such techniques as endonuclease cleavage, ligation and loopout mutagenesis.

For expression, a DNA sequence encoding tissue transglutaminase is inserted into a suitable expression vector, which in turn is used to transform or transfect appropriate host cells for expression. Expression vectors for use in carrying out the present invention will comprise a promoter capable of directing the transcription of a cloned DNA and a transcriptional terminator, operably linked with the sequence encoding the tissue transglutaminase so as to produce a

continuously transcribable gene sequence which produces sequences in reading frame and continuously translated to produce a transglutaminase polypeptide.

Host cells for use in practicing the present invention include mammalian, avian, plant, insect, bacterial and fungal cells, but preferably eukaryotic cells. Preferred eukaryotic cells include cultured mammalian cell lines (e.g., rodent or human cell lines) and fungal cells, including species of yeast (e.g., Saccharomyces spp., particularly S. cerevisiae, Schizosaccharomyces spp., or Kluyveromyces spp.) or filamentous fungi (e.g., Aspergillus spp., Neurospora spp.). Methods for producing recombinant proteins in a variety of prokaryotic and eukaryotic host cells are generally known in the art.

Techniques for transforming fungi are well known in the literature, and have been described, for instance, by Beggs (ibid.), Hinnen et al. (Proc. Natl. Acad. Sci. USA 75: 1929-1933, 1978), Yelton et al. (Proc. Natl. Acad. Sci. USA 81: 1740-1747, 1984), and Russell (Nature 301: 167-169, 1983). The genotype of the host cell will generally contain a genetic defect that is complemented by the selectable marker present on the expression vector. Choice of a particular host and selectable marker is well within the level of ordinary skill in the art.

Suitable yeast vectors for use in the present invention include YRp7 (Struhl et al., Proc. Natl. Acad. Sci. USA 76: 1035-1039, 1978), YEp13 (Broach et al., Gene 8: 121-133, 1979), POT vectors (Kawasaki et al, U.S. Patent No. 4,931,373, which is incorporated by reference herein), pJDB249 and pJDB219 (Beggs, Nature 275:104-108, 1978) and derivatives thereof. Such vectors will generally include a selectable marker, which may be one of any number of genes that exhibit a dominant phenotype for which a phenotypic assay exists to enable transformants to be selected. Preferred selectable markers are those that complement host cell auxotrophy, provide antibiotic resistance or enable a cell to utilize specific carbon sources, and include LEU2 (Broach et al., ibid.), URA3 (Botstein et al., Gene 8: 17, 1979), HIS3 (Struhl et al., ibid.) or POT1 (Kawasaki et al., ibid.). Another suitable

selectable marker is the CAT gene, which confers chloramphenicol resistance on yeast cells.

Preferred promoters for use in yeast include promoters from yeast glycolytic genes (Hitzeman et al., J. Biol. Chem. 255: 12073-12080, 1980; Alber and Kawasaki, J. Mol. Appl. Genet. 1: 419-434, 1982; Kawasaki, U.S. Patent No. 4,599,311) or alcohol dehydrogenase genes (Young et al., in Genetic Engineering of Microorganisms for Chemicals, Hollaender et al., (eds.), p. 355, Plenum, New York, 1982; Ammerer, Meth. Enzymol. 101: 192-201, 1983). In this regard, particularly preferred promoters are the TPI1 promoter (Kawasaki, U.S. Patent No. 4,599,311, 1986) and the ADH2-4^{c} promoter (Russell et al., Nature 304: 652-654, 1983; Irani and Kilgore, U.S. Patent Application Serial No. 183,130, which is incorporated herein by reference). The expression units may also include a transcriptional terminator. A preferred transcriptional terminator is the TPI1 terminator (Alber and Kawasaki, ibid.).

Additional vectors, promoters and terminators for use in expressing the transglutaminases of the invention in yeast are well known in the art and are reviewed by, for example, Emr, Meth. Enzymol. 185:231-279, (1990), incorporated herein by reference.

The transglutaminases of the invention may be expressed in Aspergillus spp. (McKnight and Upshall, described in U.S. Patent 4,935,349, which is incorporated herein by reference). Useful promoters include those derived from Aspergillus nidulans glycolytic genes, such as the ADH3 promoter (McKnight et al., EMBO J. 4:2093-2099, 1985) and the tpiA promoter. An example of a suitable terminator is the ADH3 terminator (McKnight et al., ibid.). Techniques for transforming fungi are well known in the literature, and have been described, for instance by Beggs (ibid.), Hinnen et al. (Proc. Natl. Acad. Sci. USA 75:1929-1933, 1978), Yelton et al. (Proc. Natl. Acad. Sci. USA 81:1740-1747, 1984), and Russell (Nature 301:167-169, 1983) each of which are incorporated herein by reference.

In addition to fungal cells, cultured mammalian cells may be used as host cells within the present invention.

Preferred cultured mammalian cells for use in the present invention include the COS-1 (ATCC CRL 1650) and BALB/c 3T3 (ATCC CRL 163) cell lines. In addition, a number of other mammalian cell lines may be used within the present invention, including BHK (ATCC CRL 10314), 293 (ATCC CRL 1573), Rat Hep I (ATCC CRL 1600), Rat Hep II (ATCC CRL 1548), TCMK (ATCC CCL 139), Human lung (ATCC CCL 75.1), Human hepatoma (ATCC HTB-52), Hep G2 (ATCC HB 8065), Mouse liver (ATCC CCL 29.1), NCTC 1469 (ATCC CCL 9.1) and DUKX cells (Urlaub and Chasin, Proc. Natl. Acad. Sci USA 77: 4216-4220, 1980).

Mammalian expression vectors for use in carrying out the present invention will include a promoter capable of directing the transcription of a cloned gene or cDNA. Preferred promoters include viral promoters and cellular promoters. Viral promoters include the immediate early cytomegalovirus promoter (Boshart et al., Cell 41: 521-530, 1985), the SV40 promoter (Subramani et al., Mol. Cell. Biol. 1: 854-864, 1981), and the major late promoter from Adenovirus 2 (Kaufman and Sharp, Mol. Cell. Biol. 2: 1304-1319, 1982). Cellular promoters include the mouse metallothionein-1 promoter (Palmiter et al., U.S. Patent No. 4,579,821), a mouse V_{κ} promoter (Bergman et al., Proc. Natl. Acad. Sci. USA 81: 7041-7045, 1983; Grant et al., Nuc. Acids Res. 15: 5496, 1987) and a mouse V_H promoter (Loh et al., Cell 33: 85-93, 1983). Also contained in the expression vectors is a polyadenylation signal located downstream of the coding sequence of interest. Polyadenylation signals include the early or late polyadenylation signals from SV40 (Kaufman and Sharp, ibid.), the polyadenylation signal from the Adenovirus 5 E1B region and the human growth hormone gene terminator (DeNoto et al., Nuc. Acids Res. 9: 3719-3730, 1981). Vectors can also include enhancer sequences, such as the SV40 enhancer and the mouse μ enhancer (Gillies, Cell 33: 717-728, 1983). Expression vectors may also include sequences encoding the adenovirus VA RNAs. Vectors can be obtained from commercial sources (e.g., Stratagene, La Jolla, CA).

Cloned DNA sequences may be introduced into cultured mammalian cells by, for example, calcium phosphate-mediated

transfection (Wigler et al., Cell 14: 725, 1978; Corsaro and Pearson, Somatic Cell Genetics 7: 603, 1981; Graham and Van der Eb, Virology 52: 456, 1973), electroporation (Neumann et al., EMBO J. 1: 841-845, 1982), or DEAE-dextran mediated transfection (Ausubel et al., (ed.) Current Protocols in Molecular Biology, John Wiley and Sons, Inc., NY (1987), incorporated herein by reference). To identify cells that have stably integrated the cloned DNA, a selectable marker is generally introduced into the cells along with the gene or cDNA of interest. Preferred selectable markers for use in cultured mammalian cells include genes that confer resistance to drugs, such as neomycin, hygromycin, and methotrexate. The selectable marker may be an amplifiable selectable marker. Preferred amplifiable selectable markers are the DHFR gene and the neomycin resistance gene. Selectable markers are reviewed by Thilly (Mammalian Cell Technology, Butterworth Publishers, Stoneham, MA, which is incorporated herein by reference). The choice of selectable markers is well within the level of ordinary skill in the art.

Selectable markers may be introduced into the cell on a separate vector at the same time as the transglutaminase sequence of interest, or they may be introduced on the same vector. If on the same vector, the selectable marker and the transglutaminase sequence of interest may be under the control of different promoters or the same promoter, the latter arrangement producing a dicistronic message. Constructs of this type are known in the art (for example, Levinson and Simonsen, U.S. Patent No. 4,713,339). It may also be advantageous to add additional DNA, known as "carrier DNA" to the mixture which is introduced into the cells.

Transfected mammalian cells are allowed to grow for a period of time, typically 1-2 days, to begin expressing the DNA sequence(s) of interest. Drug selection is then applied to select for growth of cells that are expressing the selectable marker in a stable fashion. For cells that have been transfected with an amplifiable selectable marker the drug concentration may be increased in a stepwise manner to select for increased copy number of the cloned sequences, thereby

increasing expression levels.

Promoters, terminators and methods for introducing expression vectors encoding transglutaminase into plant, avian and insect cells are well known in the art. The use of baculoviruses, for example, as vectors for expressing heterologous DNA sequences in insect cells has been reviewed by Atkinson et al. (Pestic. Sci. 28: 215-224,1990). The use of Agrobacterium rhizogenes as vectors for expressing genes in plant cells has been reviewed by Sinkar et al. (J. Biosci. (Banglaore) 11: 47-58, 1987).

Host cells containing DNA constructs of the present invention are then cultured to produce the transglutaminase. The cells are cultured according to standard methods in a culture medium containing nutrients required for growth of the chosen host cells. A variety of suitable media are known in the art and generally include a carbon source, a nitrogen source, essential amino acids, vitamins and minerals, as well as other components, e.g., growth factors or serum, that may be required by the particular host cells. The growth medium will generally select for cells containing the DNA construct by, for example, drug selection or deficiency in an essential nutrient which is complemented by the selectable marker on the DNA construct or co-transfected with the DNA construct.

Yeast cells, for example, are preferably cultured in a medium which comprises a nitrogen source (e.g., yeast extract), inorganic salts, vitamins and trace elements. The pH of the medium is preferably maintained at a pH greater than 2 and less than 8, preferably at pH 5-6. Methods for maintaining a stable pH include buffering and constant pH control, preferably through the addition of sodium hydroxide. Preferred buffering agents include succinic acid and Bis-Tris (Sigma Chemical Co., St. Louis, MO). Cultured mammalian cells are generally cultured in commercially available serum-containing or serum-free media. Selection of a medium appropriate for the particular cell line used is within the level of ordinary skill in the art.

In a preferred embodiment, human tissue transglutaminase is expressed in yeast as an intracellular

BNSDOCID: <WO 9212238A1 I >

product. The yeast host is a diploid strain homozygous for pep4, a mutation that reduces vacuolar protease levels, as described in Jones et al., Genetics 85:23-33 (1977), incorporated herein by reference. The strain is also homozygous for disruption of the endogenous TPI (triose phosphate isomerase) gene, thereby allowing the *S. pombe* POT1 gene to be used as a selectable marker. The vector includes the POT1 marker, a leu2-d marker and the ADH2-4c promoter. The POT1 marker in the TPI^- host allows for selection by growth in glucose. The host strain is grown in glucose-containing synthetic media with a glucose feed. An ethanol feed is then substituted for glucose to de-repress the promoter. The pH is maintained with NaOH. Other preferred means for expression are generally described in, e.g., EPO publication EP 268,772, incorporated herein by reference.

The tissue transglutaminase produced according to the present invention may be purified by affinity chromatography on an antibody column using antibodies directed against transglutaminase. Additional purification may be achieved by conventional chemical purification means, such as liquid chromatography, gradient centrifugation, and gel electrophoresis, among others. Methods of protein purification are known in the art (see generally, Scopes, R., Protein Purification, Springer-Verlag, NY (1982), which is incorporated herein by reference) and may be applied to the purification of the recombinant transglutaminase described herein. Substantially pure recombinant tissue transglutaminase of at least about 50% is preferred, at least about 70-80% more preferred, and 95-99% or more homogeneity most preferred, particularly for pharmaceutical uses. Once purified, partially or to homogeneity, as desired, the recombinant tissue transglutaminase may then be used in food preparation, protein chemistry, therapeutically, etc.

The human and murine tissue transglutaminases produced according to the present invention find a variety of uses. For example, tissue transglutaminases may be used in the preparation of food material, such as paste food, cheese, and can be added to dehydrated fish to prevent deterioration caused

by protozoans, e.g., myxamoeba. The transglutaminases can also be used in the preparation of ground meat of okiomi (Euphasia superba), by adding to dehydrated meat parts from 0.1 to 100 units, preferably about 1-40 U per gram of protein to improve meat texture and quality. Frozen granular meats can be improved by combining meat material with tissue transglutaminase of the invention at 1-500 U per gram protein, at 30-60°C for 10-120 min. to promote crosslinking between glutamine groups and lysine contained in meat preparations.

Other uses of the tissue transglutaminases described herein include the enzyme-catalyzed labeling of proteins and cell membranes (Iwanij, Eur. J. Biochem. 80:359-368 (1977), incorporated herein by reference), in the introduction of cleavable crosslinks, and in the solid phase reversible removal of specific proteins from biological systems.

The human transglutaminase of the invention also can be used therapeutically in humans. For example, the transglutaminase may be used in the repair of wounds and ulcerated lesions. As the tissue enzyme is relatively stable, active extracellularly, and binds avidly to collagen, it can be used to stabilize basement membrane structures. An appropriate endogenous substrate for the enzyme is fibronectin, which thus serves as a basis for crosslinking and stabilizing collagen/fibronectin complexes.

Transglutaminase expression can be used as a marker for screening for agonists and antagonists of cellular apoptosis. Identifying agents which inhibit the expression of transglutaminase by a cell provides a means to prevent or delay atrophic changes characteristic of many degenerative changes, particularly degenerative nerve diseases, such as Parkinson's disease and Alzheimer's disease. Inhibition of apoptosis may also enhance blood cell counts in chemotherapy patients. The tissue transglutaminase or the nucleic acids which encode the tissue transglutaminase of the invention can also be used to identify agents which induce apoptotic activity by a cell, for the control of, e.g., hyperproliferative disorders. The growth of cells such as adipocytes can be regulated with agents identified using the tissue transglutaminases provided herein

as a marker, providing a means for controlling fat depots in certain forms of obesity without the necessity for surgical intervention.

Sequences which encodes transglutaminases may be directly detected in cells with labeled synthetic oligonucleotide probes in a hybridization procedure similar to the Southern or dot blot. Also, the polymerase chain reaction (Saiki et al., Science 239:487 (1988), and U.S. Pat. No. 4,683,195) may be used to amplify DNA sequences, which are subsequently detected by their characteristic size on agarose gels, Southern blot of the gels using transglutaminase sequences or a oligonucleotide probe, or a dot blot using similar probes. The probes may comprise from about 14 nucleotides to about 25 or more nucleotides, sometimes 40 to 60 nucleotides, and in some instances a substantial portion or even the entire cDNA of a transglutaminase gene of the invention may be used. The probes are labeled with a detectable signal, such as an enzyme, biotin, a radionuclide, fluorophore, chemiluminescer, paramagnetic particle, etc.

The following examples are provided by way of illustration, not limitation.

EXAMPLE I

Cloning of Mouse and Human Tissue Transglutaminases

The following example describes the cDNA cloning and screening of mouse macrophage and heart tissue transglutaminase and human endothelial cell transglutaminase.

An oligo-dT primed cDNA library was constructed in the lambda gt11 vector with polyA$^+$ RNA purified from mouse peritoneal macrophages essentially as described by Chiocca et al., J. Biol. Chem. 263:11584-11589 (1988), incorporated herein by reference. Briefly, mice (1200) were sacrificed and their peritoneal cavities were washed with RPMI. Cells were harvested from the wash by centrifugation. The cells were plated and macrophages were allowed to attach to the dishes for 60 minutes. The dishes were then washed and the macrophages

were recovered. To induce transglutaminase expression, the macrophages were stimulated with retinoic acid (10^{-6} M) for 6 hours. RNA was then isolated from the cells, cDNA was synthesized and E. coli cells were infected with recombinant phages.

To screen the recombinant clones goat polyclonal antibodies (made against guinea pig liver transglutaminase, as described in Murtagh et al., J. Biol. Chem. 258:11074-11081 (1983), incorporated herein by reference) were used. Clones TG700 and TG1600 were identified, and [^{32}P]-labeled RNA transcripts from these clones were used to rescreen the library, as generally described in Ausubel et al., supra. Two positive clones identified in the second screening were designated TG3000 and TG3400. Since none of the clones isolated in these initial rounds of screening was a full length sequence or encoded the 5'-end of the cDNA, the macrophage library was then subjected to two more rounds of screening with short hybridization probes prepared from the 5'-ends of TG3000 and TG3400 (solid bars Fig. 1). All of these clones stopped short of the translation start site, so a mouse heart cDNA library cloned into lambda ZAP (Stratagene) was screened with an oligonucleotide hybridization probe derived from the 5'-end of clone TG7.4 (Fig. 1). A heart transglutaminase cDNA isolated by this procedure (clone TGHZ3) included 29 nucleotides of 5'-untranslated region, the initiation codon ATG and 1775 nucleotides of coding sequence.

Primer extension analysis of RNA derived from control and retinoic acid stimulated mouse macrophages, performed as generally described in Ausubel et al., ibid., and Maniatis et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1982), incorporated by reference herein, was used to locate the transcription start site. A synthetic oligonucleotide of sequences included in the 5' end of cDNA clone TG7.4 was hybridized to macrophage RNA and then transcripts were synthesized in the presence of ^{32}P-dATP. A single prominent 255 nucleotide band, more abundant in transcripts from retinoic acid-stimulated than control macrophage RNA was detected, locating the transcription start

site at a position 103 nucleotides upstream of the initiation ATG.

Repetitive screening of a mouse macrophage and heart cDNA libraries resulted in the isolation of 12 overlapping cDNA clones encoding 3487 nucleotides that spanned nearly the entire mouse tissue transglutaminase mRNA. The overlapping clones included a consensus polyadenylation signal at the 3'-end. The 5'-end of the clone isolated from mouse heart library was 75 nucleotides short of the transcript start site. The clones were sequenced as described in Example II below.

To clone the human endothelial cell transglutaminase, an oligo-dT and random-primed cDNA library was constructed with polyA$^+$ RNA from human umbilical vein endothelial cells (HUVEC) in the vector lambda ZAP (Stratagene Inc., La Jolla, CA). X-LI blue cells (Stratagene) were infected with recombinant phages and 2 x 10^5 plaques were screened with a random-primed mouse macrophage transglutaminase [^{32}P]-labeled DNA probe spanning the active site and the putative calcium binding domain (TG7.4) to facilitate isolation of full length cDNA. The hybridization procedure was done generally as described in Ausubel et al., ibid., at 55°C overnight with a final wash at 60°C in 0.1% SSC/0.1% SDS for 30 minutes. The initial screen of 2 x 10^5 recombinant phage yielded 5 positive clones. The inserts in three of the clones, hTG2, 3 and 5, were totally included within the largest cDNA clone, hTG1, which was approximately 3.3 kilobases.

To locate the transcription start site of human endothelial transglutaminase mRNA, a primer complementary to nucleotide positions 53 to 72 was synthesized and used to determine the size of the cDNA extension product. This reaction resulted in a transcript of 52 nucleotides, indicating the transcription start site is 135 nucleotides upstream from the initiator ATG and is likely coincident with the 5'-end of clone hTG-1.

EXAMPLE II

Nucleotide Sequence Analysis of Human and Mouse Tissue Transglutaminase cDNA Clones

The following Example describes the sequencing of mouse and human cDNA clones obtained in Example I. The results show a substantial degree of sequence homology between the two species of tissue enzyme.

To determine the human tissue transglutaminase sequence, both strands of the human tissue transglutaminase cDNA clone (hTG-1) were sequenced by the dideoxy chain termination method with a Se enase enzyme kit using synthetic oligonucleotide primers and deleted clones derived by exonuclease digestion (Fig. 2). As shown in Fig. 3 (SEQ. ID. NO. 1), the 3257 nucleotides included a single open reading frame encoding 687 amino acids (also SEQ.ID. NO. 2). An initiation codon, located 136-138 nucleotides downstream from the transcription start site, was included within a consensus sequence (ACCATGG) recognized as optimal for the initiation of eukaryotic translation (Kozak sequence; Kozak, Cell 44:283-292 (1986)). A terminator codon (TAA), located at nucleotide 2194-2196, was followed by 1058 nucleotides of 3'-untranslated sequence. No consensus polyadenylation signal sequence was recognized in the 3'-untranslated region.

The nucleotide sequence of the mouse tissue transglutaminase (Fig. 3) (SEQ. ID. No. 3) was determined by sequencing of overlapping cDNA clones (Fig. 1). CsCl purified, double-stranded mouse cDNA was sequenced by both the chemical degradation (Maxam and Gilbert, Proc. Natl. Acad. Sci. USA 74:560-564 (1977)) and the dideoxy chain termination (Sanger et al., Proc. Natl. Acad. Sci. USA 74:5463-5467 (1977)) methods.

The entire murine transcription unit sequence was slightly larger than the human transglutaminase cDNA clone (Fig. 3). An open reading frame of 2055 nucleotides encoded a protein of 685 amino acid residues (SEQ. ID. NO. 4). The initiation ATG was included in a Kozak sequence. The terminator triplet (TAA) was followed by 1400 bp of 3'-untranslated sequence that included a consensus polyadenylation

sequence (AATAAA) at the 3' end of the clone.

The DNA and amino acid sequences were analyzed with computer programs NUCALN and PRTALN. See Wilbur and Lipman, Proc. Natl. Acad. Sci. USA 80:726-730 (1983), incorporated herein by reference. Fig. 3 compares the nucleotide sequences of the human and mouse tissue transglutaminase cDNA's. The sequence of the human enzyme is fully represented in the top line and the deduced amino acid sequence of the coding domain is shown below. The third line contains the deduced amino acid sequence of the mouse tissue transglutaminase. Residues identical with the human enzyme are shown in an asterisk, residues that are distinct are shown with the single letter code. The fourth line shows the nucleotide sequence of the mouse transglutaminase. Nucleotides identical to the human enzyme are shown with a dash.

Comparison of the overall nucleotide sequence of the human and mouse coding domains shows a high degree (>82%) of homology. Most substitutions are silent mutations in the third position of codons. The 3' untranslated region of the two cDNA's showed no significant homology. The overall homology at the amino acid level was 84%. Table I compares the amino acid composition and the calculated molecular weight of human and mouse tissue transglutaminase.

TABLE I: Amino acid composition of human and mouse tissue transglutaminases.

amino acid	Human Endothelial	Mouse Macrophage
	n. res./mol.	
F =	25	22
L =	69	70
I =	32	30
M =	11	9
V =	57	54
S =	40	48
P =	32	30
T =	35	31
A =	40	37
Y =	23	27
H =	13	11
E =	51	52
Q =	26	25
D =	38	44
N =	34	35
K =	32	32
C =	20	20
W =	13	13
R =	39	41
G =	51	51
deduced Mr =	77253	76699

EXAMPLE V

Northern Blot Analysis of Human and Mouse Transglutaminase mRNA

Hybridization probes suitable for detecting tissue transglutaminase mRNA in tissues were prepared as TG1600 antisense RNA (^{32}P labeled using ^{32}P-UTP). Northern blots were performed according to Thomas, Proc. Natl. Acad. Sci. USA 77:5201 (1980), incorporated herein by reference. Fig. 4 shows the Northern blot analysis of RNA from human umbilical vein

endothelial cells probed with radiolabeled cDNA prepared form the insert in clone hTG-1. A single band at approximately 3.5 kilobases was detected. Fig. 5 shows the results of Northern analysis of RNA's prepared from several mouse tissues (liver, spleen, kidney, testis, heart, lungs, thymus, and brain). Minimal levels of transglutaminase mRNA were detected in thymus and in brain tissues. The levels of this RNA were higher in liver, spleen and testis and were highest in the kidney, lung and heart.

EXAMPLE VI

Expression of Tissue Transglutaminase in Eukaryotic Cells

A cDNA clone for human endothelial cell tissue transglutaminase (clone hTG1) was cloned into the Eco RI site of the eukaryotic expression vector pSG5 (Stratagene). This 3257 bp insert contained 138 nucleotides of 5'-untranslated sequence, the coding region of the enzyme and 1058 bp of 3'-untranslated sequence.

The human tissue transglutaminase expression plasmid was transiently transfected into COS-1 cells using a DEAE-Dextran mediated transfection protocol. Cells were cultured in DMEM containing 10% fetal calf serum (FCS). After 48 and 72 hours cells were washed, scraped and homogenized and the transglutaminase activity was measured as the calcium-dependent covalent incorporation of radiolabeled putrescine into N,N-dimethylcasein (essentially as described by Murtagh et al., J. Biol. Chem. 261:614-621 (1986)). In control COS-1 cells transglutaminase activity was 5.6 fmols/min/mg. In the cells transfected with the transglutaminase expression vector the transglutaminase activity was 270 fmols/min/mg.

BALB/c 3T3 cells were co-transfected (via the $CaPO_4$ procedure) with the transglutaminase expression vector and an SV-neo containing plasmid (obtained from Clontech). The cells were grown for 48 hours in DMEM containing 10% FCS and 10% Serum Plus (Hazelton Biologics, Inc., Lexena, KS). The cells

were washed and the medium was replaced with DMEM containing 10% FCS, 10% Serum Plus and 400 μg/ml G-418. G-418 resistant cells were cloned. Individual clones of transfected 3T3 cells were grown to confluency. The cells were then lysed and expression of tissue transglutaminase was measured by Western blot (U.S. Pat. No. 4,452,901; Towbin et al., Proc. Natl. Acad. Sci. USA 76:4350-4358 (1979)) using a polyclonal antibody to guinea pig liver tissue transglutaminase and by enzymatic assay (using the assay described above). Three clones of 3T3 cells (clones 13, 15, and 19) were isolated and characterized in detail. Western blot of the three transfected clones showed a prominent 80,000 kD band of immunoreactivity in the three cell lines, with the abundance in the order cl 15 >> cl 13 > cl 19. No immunoreactive tissue transglutaminase was detected in the non-transfected 3T3 cell extracts. Enzymatic assay of homogenates of the transfected cells showed activities as follows:

Control 3T3	< 0.17	pmol/min/mg
Clone 19	.051	"
Clone 13	.188	"
Clone 15	.775	"

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Davies, Peter JA
Stein, Joseph P

(ii) TITLE OF INVENTION: CLONING AND EXPRESSION OF TISSUE TRANSGLUTAMINASE

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Steven W. Parmelee
(B) STREET: One Market Plaza, Steuart Tower, Suite 2000
(C) CITY: San Francisco
(D) STATE: California
(E) COUNTRY: U.S.A.
(F) ZIP: 94105

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.24

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 07/635,756
(B) FILING DATE: 04-JAN-1991
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Parmelee, Steven W.
(B) REGISTRATION NUMBER: 31,990
(C) REFERENCE/DOCKET NUMBER: 13952-7

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (415) 543-9600
(B) TELEFAX: (415) 543-5043

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3257 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

BNSDOCID: <WO 9212238A1 I >

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens
(D) DEVELOPMENTAL STAGE: Adult
(F) TISSUE TYPE: Umbilical vein
(G) CELL TYPE: Endothelial
(H) CELL LINE: HUVEC

(vii) IMMEDIATE SOURCE:
(B) CLONE: hTG-1

(viii) POSITION IN GENOME:
(C) UNITS: bp

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 136..2199
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
AACAGGCGTG ACGCCAGTTC TAAACTTGAA ACAAAACAAA ACTTCAAAGT ACACCAAAAT    60

AGAACCTCCT TAAAGCATAA ATCTCACGGA GGGTCTCGGC CGCCAGTGGA AGGAGCCACC   120

GCCCCCGCCC CGACC ATG GCC GAG GAG CTG GTC TTA GAG AGG TGT GAT CTG    171
                 Met Ala Glu Glu Leu Val Leu Glu Arg Cys Asp Leu
                  1               5                  10

GAG CTG GAG ACC AAT GGC CGA GAC CAC CAC ACG GCC GAC CTG TGC CGG     219
Glu Leu Glu Thr Asn Gly Arg Asp His His Thr Ala Asp Leu Cys Arg
        15                  20                  25

GAG AAG CTG GTG GTG CGA CGG GGC CAG CCC TTC TGG CTG ACC CTG CAC     267
Glu Lys Leu Val Val Arg Arg Gly Gln Pro Phe Trp Leu Thr Leu His
    30                  35                  40

TTT GAG GGC CGC AAC TAC GAG GCC AGT GTA GAC AGT CTC ACC TTC AGT     315
Phe Glu Gly Arg Asn Tyr Glu Ala Ser Val Asp Ser Leu Thr Phe Ser
45                  50                  55                  60

GTC GTG ACC GGC CCA GCC CCT AGC CAG GAG GCC GGG ACC AAG GCC CGT     363
Val Val Thr Gly Pro Ala Pro Ser Gln Glu Ala Gly Thr Lys Ala Arg
                65                  70                  75

TTT CCA CTA AGA GAT GCT GTG GAG GAG GGT GAC TGG ACA GCC ACC GTG     411
Phe Pro Leu Arg Asp Ala Val Glu Glu Gly Asp Trp Thr Ala Thr Val
            80                  85                  90

GTG GAC CAG CAA GAC TGC ACC CTC TCG CTG CAG CTC ACC ACC CCG GCC     459
Val Asp Gln Gln Asp Cys Thr Leu Ser Leu Gln Leu Thr Thr Pro Ala
        95                 100                 105
```

3NSDOCID: <WO 9212238A1 I >

```
AAC GCC CCC ATC GGC CTG TAT CGC CTC AGC CTG GAG GCC TCC ACT GGC    507
Asn Ala Pro Ile Gly Leu Tyr Arg Leu Ser Leu Glu Ala Ser Thr Gly
    110                 115                 120

TAC CAG GGA TCC AGC TTT GTG CTG GGC CAC TTC ATT TTG CTC TTC AAC    555
Tyr Gln Gly Ser Ser Phe Val Leu Gly His Phe Ile Leu Leu Phe Asn
125                 130                 135                 140

GCC TGG TGC CCA GCG GAT GCT GTG TAC CTG GAC TCG GAA GAG GAG CGG    603
Ala Trp Cys Pro Ala Asp Ala Val Tyr Leu Asp Ser Glu Glu Glu Arg
                145                 150                 155

CAG GAG TAT GTC CTC ACC CAG CAG GGC TTT ATC TAC CAG GGC TCG GCC    651
Gln Glu Tyr Val Leu Thr Gln Gln Gly Phe Ile Tyr Gln Gly Ser Ala
            160                 165                 170

AAG TTC ATC AAG AAC ATA CCT TGG AAT TTT GGG CAG TTT CAA GAT GGG    699
Lys Phe Ile Lys Asn Ile Pro Trp Asn Phe Gly Gln Phe Gln Asp Gly
        175                 180                 185

ATC CTA GAC ATC TGC CTG ATC CTT CTA GAT GTC AAC CCC AAG TTC CTG    747
Ile Leu Asp Ile Cys Leu Ile Leu Leu Asp Val Asn Pro Lys Phe Leu
    190                 195                 200

AAG AAC GCC GGC CGT GAC TGC TCC CGG CGC AGC AGC CCC GTC TAC GTG    795
Lys Asn Ala Gly Arg Asp Cys Ser Arg Arg Ser Ser Pro Val Tyr Val
205                 210                 215                 220

GGC CGG GTG GGT AGT GGC ATG GTC AAC TGC AAC GAT GAC CAG GGT GTG    843
Gly Arg Val Gly Ser Gly Met Val Asn Cys Asn Asp Asp Gln Gly Val
                225                 230                 235

CTG CTG GGA CGC TGG GAC AAC AAC TAC GGG GAC GGC GTC AGC CCC ATG    891
Leu Leu Gly Arg Trp Asp Asn Asn Tyr Gly Asp Gly Val Ser Pro Met
            240                 245                 250

TCC TGG ATC GGC AGC GTG GAC ATC CTG CGG CGC TGG AAG AAC CAC GGC    939
Ser Trp Ile Gly Ser Val Asp Ile Leu Arg Arg Trp Lys Asn His Gly
        255                 260                 265

TGC CAG CGC GTC AAG TAT GGC CAG TGC TGG GTC TTC GCC GCC GTG GCC    987
Cys Gln Arg Val Lys Tyr Gly Gln Cys Trp Val Phe Ala Ala Val Ala
    270                 275                 280

TGC ACA GTG CTG AGG TGC CTA GGC ATC CCT ACC CGC GTC GTG ACC AAC   1035
Cys Thr Val Leu Arg Cys Leu Gly Ile Pro Thr Arg Val Val Thr Asn
285                 290                 295                 300

TAC AAC TCG GCC CAT GAC CAG AAC AGC AAC CTT CTC ATC GAG TAC TTC   1083
Tyr Asn Ser Ala His Asp Gln Asn Ser Asn Leu Leu Ile Glu Tyr Phe
                305                 310                 315

CGC AAT GAG TTT GGG GAG ATC CAG GGT GAC AAG AGC GAG ATG ATC TGG   1131
Arg Asn Glu Phe Gly Glu Ile Gln Gly Asp Lys Ser Glu Met Ile Trp
```

320 325 330

AAC TTC CAC TGC TGG GTG GAG TCG TGG ATG ACC AGG CCG GAC CTG CAG 1179
Asn Phe His Cys Trp Val Glu Ser Trp Met Thr Arg Pro Asp Leu Gln
335 340 345

CCG GGG TAC GAG GGC TGG CAG GCC CTG GAC CCA ACG CCC CAG GAG AAG 1227
Pro Gly Tyr Glu Gly Trp Gln Ala Leu Asp Pro Thr Pro Gln Glu Lys
350 355 360

AGC GAA GGA ACG TAC TGC TGT GGC CCA GTT CCA GTT CGT GCC ATC AAG 1275
Ser Glu Gly Thr Tyr Cys Cys Gly Pro Val Pro Val Arg Ala Ile Lys
365 370 375 380

GAG GGC GAC CTG AGC ACC AAG TAC GAT GCG CCC TTT GTC TTT GCG GAG 1323
Glu Gly Asp Leu Ser Thr Lys Tyr Asp Ala Pro Phe Val Phe Ala Glu
385 390 395

GTC AAT GCC GAC GTG GTA GAC TGG ATC CAG CAG GAC GAT GGG TCT GTG 1371
Val Asn Ala Asp Val Val Asp Trp Ile Gln Gln Asp Asp Gly Ser Val
400 405 410

CAC AAA TCC ATC AAC CGT TCC CTG ATC GTT GGG CTG AAG ATC AGC ACT 1419
His Lys Ser Ile Asn Arg Ser Leu Ile Val Gly Leu Lys Ile Ser Thr
415 420 425

AAG AGC GTG GGC CGA GAC GAG CGG GAG GAT ATC ACC CAC ACC TAC AAA 1467
Lys Ser Val Gly Arg Asp Glu Arg Glu Asp Ile Thr His Thr Tyr Lys
430 435 440

TAC CCA GAG GGG TCC TCA GAG GAG AGG GAG GCC TTC ACA AGG GCG AAC 1515
Tyr Pro Glu Gly Ser Ser Glu Glu Arg Glu Ala Phe Thr Arg Ala Asn
445 450 455 460

CAC CTG AAC AAA CTG GCC GAG AAG GAG GAG ACA GGG ATG GCC ATG CGG 1563
His Leu Asn Lys Leu Ala Glu Lys Glu Glu Thr Gly Met Ala Met Arg
465 470 475

ATC CGT GTG GGC CAG AGC ATG AAC ATG GGC AGT GAC TTT GAC GTC TTT 1611
Ile Arg Val Gly Gln Ser Met Asn Met Gly Ser Asp Phe Asp Val Phe
480 485 490

GCC CAC ATC ACC AAC AAC ACC GCT GAG GAG TAC GTC TGC CGC CTC CTG 1659
Ala His Ile Thr Asn Asn Thr Ala Glu Glu Tyr Val Cys Arg Leu Leu
495 500 505

CTC TGT GCC CGC ACC GTC AGC TAC AAT GGG ATC TTG GGG CCC GAG TGT 1707
Leu Cys Ala Arg Thr Val Ser Tyr Asn Gly Ile Leu Gly Pro Glu Cys
510 515 520

GGC ACC AAG TAC CTG CTC AAC CTA ACC CTG GAG CCT TTC TCT GAG AAG 1755
Gly Thr Lys Tyr Leu Leu Asn Leu Thr Leu Glu Pro Phe Ser Glu Lys
525 530 535 540

AGC GTT CCT CTT TGC ATC CTC TAT GAG AAA TAC CGT GAC TGC CTT ACG 1803
Ser Val Pro Leu Cys Ile Leu Tyr Glu Lys Tyr Arg Asp Cys Leu Thr

545 550 555

GAG TCC AAC CTC ATC AAG GTG CGG GCC CTC CTC GTG GAG CCA GTT ATC 1851
Glu Ser Asn Leu Ile Lys Val Arg Ala Leu Leu Val Glu Pro Val Ile
560 565 570

AAC AGC TAC CTG CTG GCT GAG AGG GAC CTC TAC CTG GAG AAT CCA GAA 1899
Asn Ser Tyr Leu Leu Ala Glu Arg Asp Leu Tyr Leu Glu Asn Pro Glu
575 580 585

ATC AAG ATC CGG ATC CTT GGG GAG CCC AAG CAG AAA CGC AAG CTG GTG 1947
Ile Lys Ile Arg Ile Leu Gly Glu Pro Lys Gln Lys Arg Lys Leu Val
590 595 600

GCT GAG GTG TCC CTG CAG AAC CCG CTC CCT GTG GCC CTG GAA GGC TGC 1995
Ala Glu Val Ser Leu Gln Asn Pro Leu Pro Val Ala Leu Glu Gly Cys
605 610 615 620

ACC TTC ACT GTG GAG GGG GCC GGC CTG ACT GAG GAG CAG AAG ACG GTG 2043
Thr Phe Thr Val Glu Gly Ala Gly Leu Thr Glu Glu Gln Lys Thr Val
625 630 635

GAG ATC CCA GAC CCC GTG GAG GCA GGG GAG GAA GTT AAG GTG AGA ATG 2091
Glu Ile Pro Asp Pro Val Glu Ala Gly Glu Glu Val Lys Val Arg Met
640 645 650

GAC CTC GTG CCG CTC CAC ATG GGC CTC CAC AAG CTG GTG GTG AAC TTC 2139
Asp Leu Val Pro Leu His Met Gly Leu His Lys Leu Val Val Asn Phe
655 660 665

GAG AGC GAC AAG CTG AAG GCT GTG AAG GGC TTC CGG AAT GTC ATC ATT 2187
Glu Ser Asp Lys Leu Lys Ala Val Lys Gly Phe Arg Asn Val Ile Ile
670 675 680

GGC CCC GCC TAA GGGACCCCTG CTCCCAGCCT GCTGAGAGCC CCCACCTTGA 2239
Gly Pro Ala
685

TCCCAATCCT TATCCCAAGC TAGTGAGCAA AATATGCCCC TTATTGGGCC CCAGACCCCA 2299

GGGCAGGGTG GGCAGCCTAT GGGGGCTCTC GGAAATGGAA TGTGCCCCTC GCCCATCTCA 2359

GCCTCCTGAG CCTGTGGGTC CCCACTCACC CCCTTTGCTG TGAGGAATGC TCTGTGCCAG 2419

AAACAGTGGG AGCCCTGACC TGTGCTGACT GGGGCTGGGG TGAGAGAGGA AAGACCTACA 2479

TTCCCTCTCC TGCCCAGATG CCCTTTGGAA AGCCATTGAC CACCCACCAT ATTGTTTGAT 2539

CTACTTCATA GCTCCTTGGA GCAGGCAAAA AAGGACAGC ATGCCCTTGG CTGGATCAGG 2599

AATCCAGCTC CCTAGACTGC ATCCCGTACC TCTTCCCATG ACTGCACCCA GCTCCAGGGG 2659

CCCTTGGGAC ACCCAGAGCT GGGTGGGGAC AGTGATAGGC CCAAGGTCCC CTCCACATCC 2719

CAGCAGCCCA AGCTTAATAG CCCTCCCCCT CAACCTCACC ATTGTGAAGC ACCTACTATG 2779

TGCTGGGTGC CTCCCACACT TGCTGGGGCT CACGGGGCCT CCAACCCATT TAATCACCAT 2839

GGGAAACTGT TGTGGGCGCT GCTTCCAGGA TAAGGAGACT GAGGCTTAGA GAGAGGAGGC 2899

AGCCCCCTCC ACACCAGTGG CCTCGTGGTT ATAAGCAAGG CTGGGTAATG TGAAGGCCCA 2959

AGAGCAGAGT CTGGGCCTCT GACTCTGAGT CCACTGCTCC ATTTATAACC CCAGCCTGAC 3019

CTGAGACTGT CGCAGAGGCT GTCTGGGGCC TTTATCAAAA AAAGACTCAG CCAAGACAAG 3079

GAGGTAGAGA GGGGACTGGG GGACTGGGAG TCAGAGCCCT GGCTGGGTTC AGGTCCCACG 3139

TCTGGCCAGG CACTGCCTTC TCCTCTCTGG GCCTTTGTTT CCTTGTTGGT CAGAGGAGTG 3199

ATTGAACCTG CTCATCTCCA AGGATCCTCT CCACTCCATG TTTGCAATAC ACAATTCC 3257

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 687 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Glu Glu Leu Val Leu Glu Arg Cys Asp Leu Glu Leu Glu Thr
1 5 10 15

Asn Gly Arg Asp His His Thr Ala Asp Leu Cys Arg Glu Lys Leu Val
20 25 30

Val Arg Arg Gly Gln Pro Phe Trp Leu Thr Leu His Phe Glu Gly Arg
35 40 45

Asn Tyr Glu Ala Ser Val Asp Ser Leu Thr Phe Ser Val Val Thr Gly
50 55 60

Pro Ala Pro Ser Gln Glu Ala Gly Thr Lys Ala Arg Phe Pro Leu Arg
65 70 75 80

Asp Ala Val Glu Glu Gly Asp Trp Thr Ala Thr Val Val Asp Gln Gln
85 90 95

Asp Cys Thr Leu Ser Leu Gln Leu Thr Thr Pro Ala Asn Ala Pro Ile
100 105 110

Gly Leu Tyr Arg Leu Ser Leu Glu Ala Ser Thr Gly Tyr Gln Gly Ser
115 120 125

Ser Phe Val Leu Gly His Phe Ile Leu Leu Phe Asn Ala Trp Cys Pro
130 135 140

Ala Asp Ala Val Tyr Leu Asp Ser Glu Glu Glu Arg Gln Glu Tyr Val
145 150 155 160

Leu Thr Gln Gln Gly Phe Ile Tyr Gln Gly Ser Ala Lys Phe Ile Lys
165 170 175

Asn Ile Pro Trp Asn Phe Gly Gln Phe Gln Asp Gly Ile Leu Asp Ile
180 185 190

Cys Leu Ile Leu Leu Asp Val Asn Pro Lys Phe Leu Lys Asn Ala Gly
195 200 205

Arg Asp Cys Ser Arg Arg Ser Ser Pro Val Tyr Val Gly Arg Val Gly
210 215 220

Ser Gly Met Val Asn Cys Asn Asp Asp Gln Gly Val Leu Leu Gly Arg
225 230 235 240

Trp Asp Asn Asn Tyr Gly Asp Gly Val Ser Pro Met Ser Trp Ile Gly
245 250 255

Ser Val Asp Ile Leu Arg Arg Trp Lys Asn His Gly Cys Gln Arg Val
260 265 270

Lys Tyr Gly Gln Cys Trp Val Phe Ala Ala Val Ala Cys Thr Val Leu
275 280 285

Arg Cys Leu Gly Ile Pro Thr Arg Val Val Thr Asn Tyr Asn Ser Ala
290 295 300

His Asp Gln Asn Ser Asn Leu Leu Ile Glu Tyr Phe Arg Asn Glu Phe
305 310 315 320

Gly Glu Ile Gln Gly Asp Lys Ser Glu Met Ile Trp Asn Phe His Cys
325 330 335

Trp Val Glu Ser Trp Met Thr Arg Pro Asp Leu Gln Pro Gly Tyr Glu
340 345 350

Gly Trp Gln Ala Leu Asp Pro Thr Pro Gln Glu Lys Ser Glu Gly Thr
355 360 365

Tyr Cys Cys Gly Pro Val Pro Val Arg Ala Ile Lys Glu Gly Asp Leu
370 375 380

Ser Thr Lys Tyr Asp Ala Pro Phe Val Phe Ala Glu Val Asn Ala Asp
385 390 395 400

Val Val Asp Trp Ile Gln Gln Asp Asp Gly Ser Val His Lys Ser Ile
405 410 415

Asn Arg Ser Leu Ile Val Gly Leu Lys Ile Ser Thr Lys Ser Val Gly
420 425 430

Arg Asp Glu Arg Glu Asp Ile Thr His Thr Tyr Lys Tyr Pro Glu Gly
435 440 445

INSDOCID: <WO 9212238A1 I >

```
Ser Ser Glu Glu Arg Glu Ala Phe Thr Arg Ala Asn His Leu Asn Lys
    450                 455                 460

Leu Ala Glu Lys Glu Glu Thr Gly Met Ala Met Arg Ile Arg Val Gly
465                 470                 475                 480

Gln Ser Met Asn Met Gly Ser Asp Phe Asp Val Phe Ala His Ile Thr
                485                 490                 495

Asn Asn Thr Ala Glu Glu Tyr Val Cys Arg Leu Leu Leu Cys Ala Arg
            500                 505                 510

Thr Val Ser Tyr Asn Gly Ile Leu Gly Pro Glu Cys Gly Thr Lys Tyr
        515                 520                 525

Leu Leu Asn Leu Thr Leu Glu Pro Phe Ser Glu Lys Ser Val Pro Leu
    530                 535                 540

Cys Ile Leu Tyr Glu Lys Tyr Arg Asp Cys Leu Thr Glu Ser Asn Leu
545                 550                 555                 560

Ile Lys Val Arg Ala Leu Leu Val Glu Pro Val Ile Asn Ser Tyr Leu
                565                 570                 575

Leu Ala Glu Arg Asp Leu Tyr Leu Glu Asn Pro Glu Ile Lys Ile Arg
            580                 585                 590

Ile Leu Gly Glu Pro Lys Gln Lys Arg Lys Leu Val Ala Glu Val Ser
        595                 600                 605

Leu Gln Asn Pro Leu Pro Val Ala Leu Glu Gly Cys Thr Phe Thr Val
    610                 615                 620

Glu Gly Ala Gly Leu Thr Glu Glu Gln Lys Thr Val Glu Ile Pro Asp
625                 630                 635                 640

Pro Val Glu Ala Gly Glu Glu Val Lys Val Arg Met Asp Leu Val Pro
                645                 650                 655

Leu His Met Gly Leu His Lys Leu Val Val Asn Phe Glu Ser Asp Lys
            660                 665                 670

Leu Lys Ala Val Lys Gly Phe Arg Asn Val Ile Ile Gly Pro Ala
        675                 680                 685
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3486 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

BNSDOCID: <WO 9212238A1 I >

(iii) HYPOTHETICAL: N

(iv) ANTI-SENSE: N

(vi) ORIGINAL SOURCE:
(A) ORGANISM: mouse
(D) DEVELOPMENTAL STAGE: Adult
(F) TISSUE TYPE: peritoneal
(G) CELL TYPE: macrophage

(viii) POSITION IN GENOME:
(C) UNITS: bp

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 30..2087
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
CGTCTGAGCT GTCGCCGCTA GCCTGGCCC ATG GCA GAG GAG CTG CTC CTA GAG      53
                                Met Ala Glu Glu Leu Leu Leu Glu
                                 1               5

AGG TGT GAT CTG GAG ATT CAG GCC AAT GGC CGT GAC CAC CAC ACG GCC     101
Arg Cys Asp Leu Glu Ile Gln Ala Asn Gly Arg Asp His His Thr Ala
    10                  15                  20

GAC CTA TGC CAA GAG AAA CTG CTG GTG CGT CGT GGC CAG CGC TTC CGG     149
Asp Leu Cys Gln Glu Lys Leu Leu Val Arg Arg Gly Gln Arg Phe Arg
25                  30                  35                  40

CTG ACT CTG TAC TTC GAG GGC CGT GGC TAC GAG GCC AGC GTG GAC AGC     197
Leu Thr Leu Tyr Phe Glu Gly Arg Gly Tyr Glu Ala Ser Val Asp Ser
                45                  50                  55

CTC ACG TTC GGT GCT GTG ACC GGC CCA GAT CCC AGT GAG GAG GCA GGG     245
Leu Thr Phe Gly Ala Val Thr Gly Pro Asp Pro Ser Glu Glu Ala Gly
            60                  65                  70

ACC AAG GCC CGC TTT TCA CTG TCT GAC AAT GTG GAG GAG GGA TCT TGG     293
Thr Lys Ala Arg Phe Ser Leu Ser Asp Asn Val Glu Glu Gly Ser Trp
        75                  80                  85

TCA GCC TCA CTG GTG GAC CAG CAG GAC AAT GTC CTG TCG CTG CAG CTC     341
Ser Ala Ser Leu Val Asp Gln Gln Asp Asn Val Leu Ser Leu Gln Leu
    90                  95                  100

TGC ACC CCA GCC AAT GCT CCT ATT GGC CTG TAC CGT CTC AGC CTA GAG     389
Cys Thr Pro Ala Asn Ala Pro Ile Gly Leu Tyr Arg Leu Ser Leu Glu
105                 110                 115                 120

GCT TCT ACT GGC TAC CAG GGC TCC AGC TTT GTG CTG GGC CAC TTC ATC     437
Ala Ser Thr Gly Tyr Gln Gly Ser Ser Phe Val Leu Gly His Phe Ile
```

125 130 135

TTG CTC TAC AAT GCC TGG TGC CCA GCC GAT GAT GTG TAC CTA GAC TCA 485
Leu Leu Tyr Asn Ala Trp Cys Pro Ala Asp Asp Val Tyr Leu Asp Ser
140 145 150

GAG GAG GAG CGA CGG GAA TAT GTC CTT ACG CAA CAG GGC TTC ATC TAC 533
Glu Glu Glu Arg Arg Glu Tyr Val Leu Thr Gln Gln Gly Phe Ile Tyr
155 160 165

CAA GGC TCT GTC AAG TTC ATC AAG AGT GTG CCT TGG AAC TTT GGG CAG 581
Gln Gly Ser Val Lys Phe Ile Lys Ser Val Pro Trp Asn Phe Gly Gln
170 175 180

TTC CAA GAT GGA ATC CTG GAC ACC TGC CTG ATG CTG TTG GAT ATG AAC 629
Phe Gln Asp Gly Ile Leu Asp Thr Cys Leu Met Leu Leu Asp Met Asn
185 190 195 200

CCC AAG TTC CTG AAG AAC CGT AGT CGG GAC TGC TCA CGC CGC AGC AGT 677
Pro Lys Phe Leu Lys Asn Arg Ser Arg Asp Cys Ser Arg Arg Ser Ser
205 210 215

CCC ATC TAT GTG GGC CGC GTG GTG AGC GAC ATG GTC AAC TGC AAT GAT 725
Pro Ile Tyr Val Gly Arg Val Val Ser Asp Met Val Asn Cys Asn Asp
220 225 230

GAC CAG GGT GTG CTT CTG GGC CGC TGG GAC AAC AAC TAT GGG GAT GGT 773
Asp Gln Gly Val Leu Leu Gly Arg Trp Asp Asn Asn Tyr Gly Asp Gly
235 240 245

ATC AGT CCC ATG GCC TGG ATT GGC AGT GTG GAC ATT CTG CGG CGT TGG 821
Ile Ser Pro Met Ala Trp Ile Gly Ser Val Asp Ile Leu Arg Arg Trp
250 255 260

AAG GAA CAC GGC TGT CAG CAA GTC AAG TAC GGC CAG TGC TGG GTG TTT 869
Lys Glu His Gly Cys Gln Gln Val Lys Tyr Gly Gln Cys Trp Val Phe
265 270 275 280

GCA GCG GTG GCC TGC ACA GTG CTG CGG TGC CTC GGC ATC CAT AAC CGG 917
Ala Ala Val Ala Cys Thr Val Leu Arg Cys Leu Gly Ile His Asn Arg
285 290 295

GTC GTG ACC AAC TAC AAC TCC GCC CAC GAC CAG AAC AGC AAC CTG CTC 965
Val Val Thr Asn Tyr Asn Ser Ala His Asp Gln Asn Ser Asn Leu Leu
300 305 310

ATC GAG TAC TTC CGA AAT GAG TTT GGG GAG CTG GAG ACG AAC AAG AGC 1013
Ile Glu Tyr Phe Arg Asn Glu Phe Gly Glu Leu Glu Thr Asn Lys Ser
315 320 325

GAG ATG ATC TGG AAC TTC CAC TGC TGG GTG GAG TCC TGG ATG ACC AGG 1061
Glu Met Ile Trp Asn Phe His Cys Trp Val Glu Ser Trp Met Thr Arg
330 335 340

CCA GAC CTA CAG CCG GGA TAT GAG GGC TGG GAG GCC CTA GAC CCC ACA 1109
Pro Asp Leu Gln Pro Gly Tyr Glu Gly Trp Glu Ala Leu Asp Pro Thr

345 350 355 360

CCA CAG GAG AAG AGC GAA GGG ACA TAC TGT TGT GGC CCA GTC TCA GTG 1157
Pro Gln Glu Lys Ser Glu Gly Thr Tyr Cys Cys Gly Pro Val Ser Val
365 370 375

CGA GCC ATC AAG GAG GGA GAC CTG AGT ACC AAG TAT GAT GCA CCC TTC 1205
Arg Ala Ile Lys Glu Gly Asp Leu Ser Thr Lys Tyr Asp Ala Pro Phe
380 385 390

GTG TTT GCC GAG GTC AAC GCT GAT GTG GTG GAC TGG ATC CGG CAG GAC 1253
Val Phe Ala Glu Val Asn Ala Asp Val Val Asp Trp Ile Arg Gln Asp
395 400 405

GAA GGG TCT GTG CTC AAA TGG ATG AAC CGT TCC TTG GTC GTG GGG CAG 1301
Glu Gly Ser Val Leu Lys Trp Met Asn Arg Ser Leu Val Val Gly Gln
410 415 420

AAG ATC AGC ACT AAG AGT GTG GGC CGT GAT GAC CGG GAG GAC ATC ACC 1349
Lys Ile Ser Thr Lys Ser Val Gly Arg Asp Asp Arg Glu Asp Ile Thr
425 430 435 440

CAT ACA TAC AAG TAC CCA GAG GGG TCA CCC GAG GAG AGG GAA GTC TTC 1397
His Thr Tyr Lys Tyr Pro Glu Gly Ser Pro Glu Glu Arg Glu Val Phe
445 450 455

ACC AAG GCC AAC CAC CTG AAC AAA CTG GAC GAG AAA GAG GGG ACA GGG 1445
Thr Lys Ala Asn His Leu Asn Lys Leu Asp Glu Lys Glu Gly Thr Gly
460 465 470

ATG GCC ATG CGC ATC CGA GTG GGG CAG TAT GAG CAT GGC AAC GAC TTC 1493
Met Ala Met Arg Ile Arg Val Gly Gln Tyr Glu His Gly Asn Asp Phe
475 480 485

GAC GTG TTT GCC CAC ATC GGC AAC GAC ACC TCG GAG ACT CGA GAG TGT 1541
Asp Val Phe Ala His Ile Gly Asn Asp Thr Ser Glu Thr Arg Glu Cys
490 495 500

CGT CTC CTG CTC TGT GCC CGC ACT GTC AGC TAC AAC GGG GTG CTG GGG 1589
Arg Leu Leu Leu Cys Ala Arg Thr Val Ser Tyr Asn Gly Val Leu Gly
505 510 515 520

CCC GAG TGT GGC ACT GAG GAC ATC AAC CTG ACC CTG GAT CCC TAC TCT 1637
Pro Glu Cys Gly Thr Glu Asp Ile Asn Leu Thr Leu Asp Pro Tyr Ser
525 530 535

GAG AAC AGC ATC CCA CTT CGA ATC CTC TAC GAG AAG TAC AGC GGG TGC 1685
Glu Asn Ser Ile Pro Leu Arg Ile Leu Tyr Glu Lys Tyr Ser Gly Cys
540 545 550

CGT ACA GAG TCA AAC CTC ATC AAG GTG CGG GGC CTT CTC ATC GAA CCA 1733
Arg Thr Glu Ser Asn Leu Ile Lys Val Arg Gly Leu Leu Ile Glu Pro
555 560 565

GCT GCC AAC AGC TAC CTG CTG GCT GAG AGA GAT CTC TAC GTG GAG AAT 1781

Ala Ala Asn Ser Tyr Leu Leu Ala Glu Arg Asp Leu Tyr Val Glu Asn
570 575 580

CCC GAA ATC AAG ATC CGG GTT TTG GGA GAA CCC AAG CAA AAC CGC AAA 1829
Pro Glu Ile Lys Ile Arg Val Leu Gly Glu Pro Lys Gln Asn Arg Lys
585 590 595 600

CTG GTG GCT GAG GTG TCC CTG AAG AAC CCA CTT TCC GAT CCC CTG TAT 1877
Leu Val Ala Glu Val Ser Leu Lys Asn Pro Leu Ser Asp Pro Leu Tyr
605 610 615

GAC TGC ATC TTC ACT GTG GAG GGG GCT GGC CTG ACC AAG GAG CAG AAG 1925
Asp Cys Ile Phe Thr Val Glu Gly Ala Gly Leu Thr Lys Glu Gln Lys
620 625 630

TCT GTG GAA GTC TCA GAC CCG GTG CCA GCG GGC GAT TTG GTC AAG GCA 1973
Ser Val Glu Val Ser Asp Pro Val Pro Ala Gly Asp Leu Val Lys Ala
635 640 645

CGG GTC GAC CTG TCC CCG ACT GAT ATT GGC CTC CAC AAG CTG GTG GTG 2021
Arg Val Asp Leu Ser Pro Thr Asp Ile Gly Leu His Lys Leu Val Val
650 655 660

AAC TTC CAG TGT GAC AAG CTG AAG TCG GTG AAG GGT TAC CGG AAT GTT 2069
Asn Phe Gln Cys Asp Lys Leu Lys Ser Val Lys Gly Tyr Arg Asn Val
665 670 675 680

ATC ATC GGC CCG GCC TAA GGGACCCCTT CCCAGACTCA ACCCCACCAC 2117
Ile Ile Gly Pro Ala
685

CTGCCAACCC CCATTCAACC TGGTCTTTAT CCTAAGATAA TGAGCAACTT CACCCCATTC 2177

AGGCTGACAT GGCTGCCTGG GGCCTCTTCA GAAGACAGTG TACTTCTGGC CCAATCCTGT 2237

TCCTCTGGAT CTATTTCCCC ATCCTGTTCC CTTAGTGTGC ACGGAAGGTC CTGTGCCGAC 2297

ACAGTGGGTA CCTGTGGAAA GGGTAAGAGG AGAGCCATCA CCAGCACTCT GTATCTCTGC 2357

ATTGTTTGAA CTGTCTCTGG AGCCTCAGCG CAAGCACAAA GGGACCGTGC GCATGGCACC 2417

ATCGAAGGAA ACGATCCTTG GAGCAGGAAC GCTGTCGGCA CCATTTGCGC TCCTGAATGG 2477

AACCATATGT GCATGGTACT TAGATCTTAC GGTACACCAG CTAGCGTACA TCCGTGTAAC 2537

TTCAGGTGGT ACAAACTGAG GCTGCTGTGC TGTACTGGGA CAGTAGGCAG GCCATCACTT 2597

GCAGGGCCAG TGGGTGGAGC TGGAATACAG GGAATCCATC TGTGACACCA GCTCTGACCT 2657

GAGCGGGTCA GAGAGGCTAT CTGGGGATGA TGAGCCTCGC GGTGGTTAGG GTGAGGAGTT 2717

GAGGGTGGGG CGGGGAGGCC AGAGAACTGG GAGTCAGAGC TTGTGTTTAA GCCCCAAGGA 2777

GAGCTACACT CTATCCTCCT CTTCTGGGCT TGATCATTCT TCACCAGGGC AGTGCTTAAA 2837

BNSDOCID: <WO 9212238A1 I >

CCAGGGGTCT CCTGGCCCTA TACTTCTAAA GCCGTGCTTG CCAGGCTCAG TGCACTTCTG 2897

GGAAGGAGAC ACTGCACCAG ACCTGTACCC GAGGTGTCCT GCTTCTCACC TATGACTGGG 2957

TTCCTCAGAT GGCAGTTCCA GGGACTATCC ACAAGCTACT CACACAGTGC CTAGACTAGA 3017

TTTCACAGAA CGAGCCTTAA ATATCCACGC TGGTCGCTAA TCAGCAGGCC ACACCTCCAC 3077

CAGCCCCTCC TCCTCAGCCA GCAGCCTCTA GACACAGCAC GTGTCTCCAG AAAAGCGTGG 3137

AAGCCTGTGT GTGGGTCATC ATCCCAAGTT AGCCCACTCC TACCTCTCTG TGGCTGGCCC 3197

TCAACCTAGG AAGGCTGGCA GTGGTGGCCG GGTCTCTGGG CGATGAGGTC AGAGTCTCTG 3257

GATCCCCTGA AATCCCGGAG AAGAGCCTGG GAAGAATCAA ACTGATGCAT TTAACGCGTT 3317

CTGCTTTACA CAGAGGATCG CACCGTGAGC CGTGCTATCT GTCCTGTCCC CACACGGTTC 3377

TGTTTCTTTT GGTCTGTGCA AGCTTCAGGT TCCAAGCAGG CAGTAGGCAT GCTTTCAAAG 3437

CACATGTGAA CACTGAAATA AAGGTCTATT TTTCACATTC ACGGAATTC 3486

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 685 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ala Glu Glu Leu Leu Leu Glu Arg Cys Asp Leu Glu Ile Gln Ala
1 5 10 15

Asn Gly Arg Asp His His Thr Ala Asp Leu Cys Gln Glu Lys Leu Leu
20 25 30

Val Arg Arg Gly Gln Arg Phe Arg Leu Thr Leu Tyr Phe Glu Gly Arg
35 40 45

Gly Tyr Glu Ala Ser Val Asp Ser Leu Thr Phe Gly Ala Val Thr Gly
50 55 60

Pro Asp Pro Ser Glu Glu Ala Gly Thr Lys Ala Arg Phe Ser Leu Ser
65 70 75 80

Asp Asn Val Glu Glu Gly Ser Trp Ser Ala Ser Leu Val Asp Gln Gln
85 90 95

Asp Asn Val Leu Ser Leu Gln Leu Cys Thr Pro Ala Asn Ala Pro Ile
100 105 110

Gly Leu Tyr Arg Leu Ser Leu Glu Ala Ser Thr Gly Tyr Gln Gly Ser

```
        115                 120                 125

Ser Phe Val Leu Gly His Phe Ile Leu Leu Tyr Asn Ala Trp Cys Pro
    130                 135                 140

Ala Asp Asp Val Tyr Leu Asp Ser Glu Glu Glu Arg Arg Glu Tyr Val
145                 150                 155                 160

Leu Thr Gln Gln Gly Phe Ile Tyr Gln Gly Ser Val Lys Phe Ile Lys
                165                 170                 175

Ser Val Pro Trp Asn Phe Gly Gln Phe Gln Asp Gly Ile Leu Asp Thr
            180                 185                 190

Cys Leu Met Leu Leu Asp Met Asn Pro Lys Phe Leu Lys Asn Arg Ser
        195                 200                 205

Arg Asp Cys Ser Arg Arg Ser Ser Pro Ile Tyr Val Gly Arg Val Val
    210                 215                 220

Ser Asp Met Val Asn Cys Asn Asp Asp Gln Gly Val Leu Leu Gly Arg
225                 230                 235                 240

Trp Asp Asn Asn Tyr Gly Asp Gly Ile Ser Pro Met Ala Trp Ile Gly
                245                 250                 255

Ser Val Asp Ile Leu Arg Arg Trp Lys Glu His Gly Cys Gln Gln Val
            260                 265                 270

Lys Tyr Gly Gln Cys Trp Val Phe Ala Ala Val Ala Cys Thr Val Leu
        275                 280                 285

Arg Cys Leu Gly Ile His Asn Arg Val Val Thr Asn Tyr Asn Ser Ala
    290                 295                 300

His Asp Gln Asn Ser Asn Leu Leu Ile Glu Tyr Phe Arg Asn Glu Phe
305                 310                 315                 320

Gly Glu Leu Glu Thr Asn Lys Ser Glu Met Ile Trp Asn Phe His Cys
                325                 330                 335

Trp Val Glu Ser Trp Met Thr Arg Pro Asp Leu Gln Pro Gly Tyr Glu
            340                 345                 350

Gly Trp Glu Ala Leu Asp Pro Thr Pro Gln Glu Lys Ser Glu Gly Thr
        355                 360                 365

Tyr Cys Cys Gly Pro Val Ser Val Arg Ala Ile Lys Glu Gly Asp Leu
    370                 375                 380

Ser Thr Lys Tyr Asp Ala Pro Phe Val Phe Ala Glu Val Asn Ala Asp
385                 390                 395                 400

Val Val Asp Trp Ile Arg Gln Asp Glu Gly Ser Val Leu Lys Trp Met
                405                 410                 415
```

BNSDOCID: <WO 9212238A1 I >

Asn Arg Ser Leu Val Val Gly Gln Lys Ile Ser Thr Lys Ser Val Gly
420 425 430

Arg Asp Asp Arg Glu Asp Ile Thr His Thr Tyr Lys Tyr Pro Glu Gly
435 440 445

Ser Pro Glu Glu Arg Glu Val Phe Thr Lys Ala Asn His Leu Asn Lys
450 455 460

Leu Asp Glu Lys Glu Gly Thr Gly Met Ala Met Arg Ile Arg Val Gly
465 470 475 480

Gln Tyr Glu His Gly Asn Asp Phe Asp Val Phe Ala His Ile Gly Asn
485 490 495

Asp Thr Ser Glu Thr Arg Glu Cys Arg Leu Leu Leu Cys Ala Arg Thr
500 505 510

Val Ser Tyr Asn Gly Val Leu Gly Pro Glu Cys Gly Thr Glu Asp Ile
515 520 525

Asn Leu Thr Leu Asp Pro Tyr Ser Glu Asn Ser Ile Pro Leu Arg Ile
530 535 540

Leu Tyr Glu Lys Tyr Ser Gly Cys Arg Thr Glu Ser Asn Leu Ile Lys
545 550 555 560

Val Arg Gly Leu Leu Ile Glu Pro Ala Ala Asn Ser Tyr Leu Leu Ala
565 570 575

Glu Arg Asp Leu Tyr Val Glu Asn Pro Glu Ile Lys Ile Arg Val Leu
580 585 590

Gly Glu Pro Lys Gln Asn Arg Lys Leu Val Ala Glu Val Ser Leu Lys
595 600 605

Asn Pro Leu Ser Asp Pro Leu Tyr Asp Cys Ile Phe Thr Val Glu Gly
610 615 620

Ala Gly Leu Thr Lys Glu Gln Lys Ser Val Glu Val Ser Asp Pro Val
625 630 635 640

Pro Ala Gly Asp Leu Val Lys Ala Arg Val Asp Leu Ser Pro Thr Asp
645 650 655

Ile Gly Leu His Lys Leu Val Val Asn Phe Gln Cys Asp Lys Leu Lys
660 665 670

Ser Val Lys Gly Tyr Arg Asn Val Ile Ile Gly Pro Ala
675 680 685

From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

WHAT IS CLAIMED IS:

1. An isolated polynucleotide sequence which codes for human tissue transglutaminase.

2. The polynucleotide sequence of claim 1, wherein the transglutaminase is of endothelial cell origin.

3. The polynucleotide molecule of claim 1, wherein the polypeptide encoded thereby catalyzes the Ca^{++} dependent crosslinking of protein-bound glutamine and lysine residues.

4. The polynucleotide of claim 1, wherein the sequence is substantially the human transglutaminase of Fig. 3.

5. The polynucleotide of claim 1, which is a cDNA sequence.

6. An isolated polynucleotide sequence which codes for mouse tissue transglutaminase.

7. The polynucleotide sequence of claim 6, wherein the mouse transglutaminase is of macrophage origin.

8. The polynucleotide of claim 6, wherein the sequence is substantially the murine transglutaminase of Fig. 3.

9. The polynucleotide molecule of claim 6, wherein the polypeptide encoded thereby catalyzes the Ca^{++} dependent crosslinking of protein-bound glutamine and lysine residues.

10. A DNA construct for the expression of human tissue transglutaminase, which comprises the following operably linked elements:

a transcriptional promoter;

a DNA sequence encoding a human tissue transglutaminase polypeptide; and

a transcriptional terminator.

11. The polypeptide which is encoded by the DNA construct of claim 10.

12. The polypeptide of claim 11, which catalyzes Ca^{++} dependent crosslinking of protein-bound glutamine and lysine residues.

13. The polypeptide of claim 12, which has substantially the amino acid sequence of human tissue transglutaminase of Fig. 3.

14. A cultured cell transformed or transfected with the DNA construct of claim 10.

15. The cultured cell of claim 14, which is a eukaryotic cell.

16. The eukaryotic cell of claim 15, which is a yeast cell or mammalian cell.

17. A method for producing human tissue transglutaminase, which comprises cultivating eukaryotic cells transformed or transfected with the DNA construct of claim 11, and isolating the transglutaminase from the cells.

18. The method of claim 17, wherein the transformed eukaryotic cells are yeast cells.

19. A probe which comprises an oligonucleotide of at least about 14 nucleotides capable of specifically hybridizing with a gene which encodes a human or murine tissue transglutaminase polypeptide, wherein said probe is at least 85% homologous to a sequence of the human or murine transglutaminase of Fig. 3.

20. The probe of claim 19, which is labeled to provide a detectable signal.

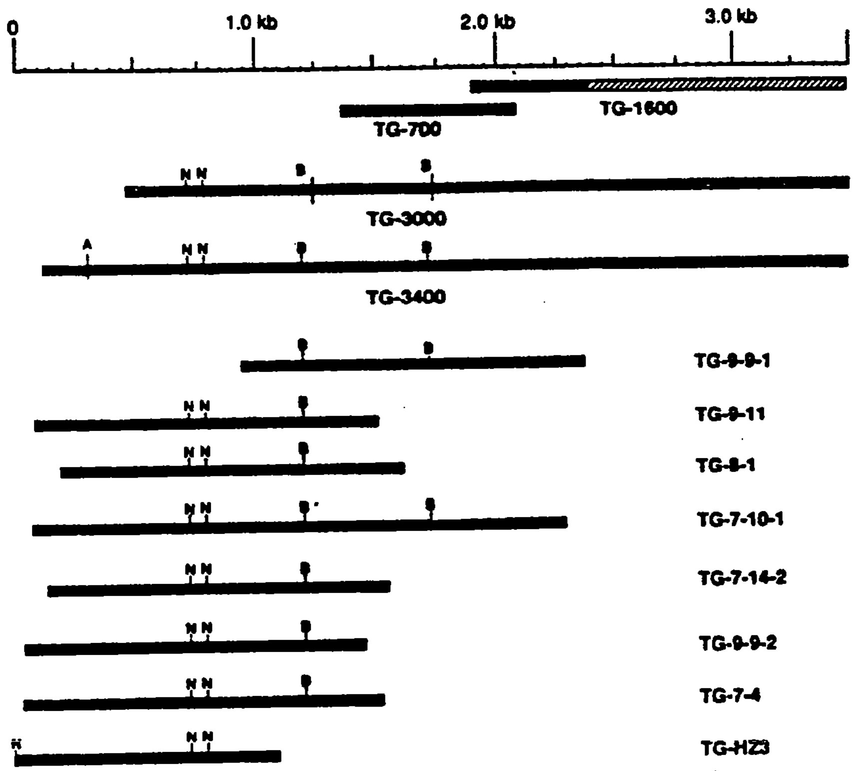

0
1.0 kb
2.0 kb
3.0 kb
TG-1600
TG-700
TG-3000
TG-3400
TG-9-9-1
TG-9-11
TG-8-1
TG-7-10-1
TG-7-14-2
TG-9-9-2
TG-7-4
TG-HZ3

Fig. 1

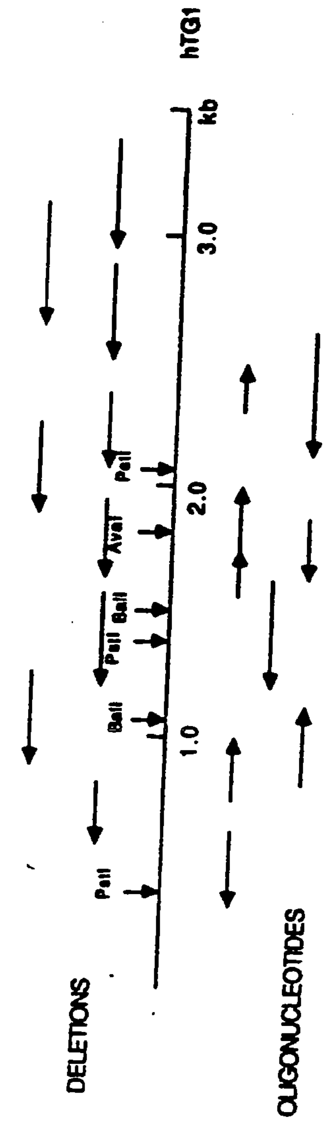


Fig. 2

BNSDOCID: <WO 9212238A1 I >

Fig. 3

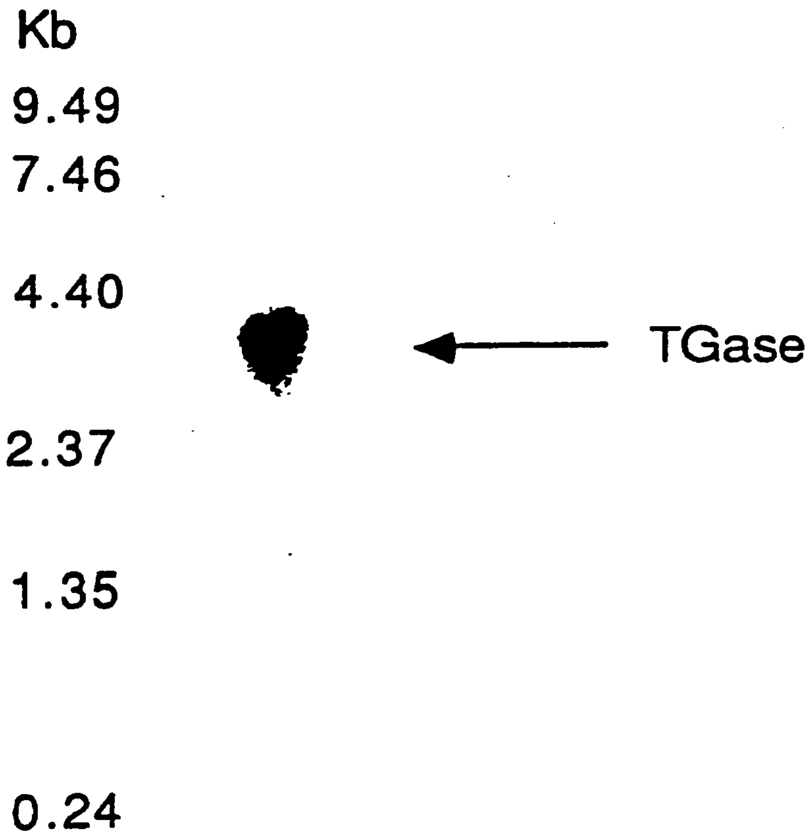


Fig. 4

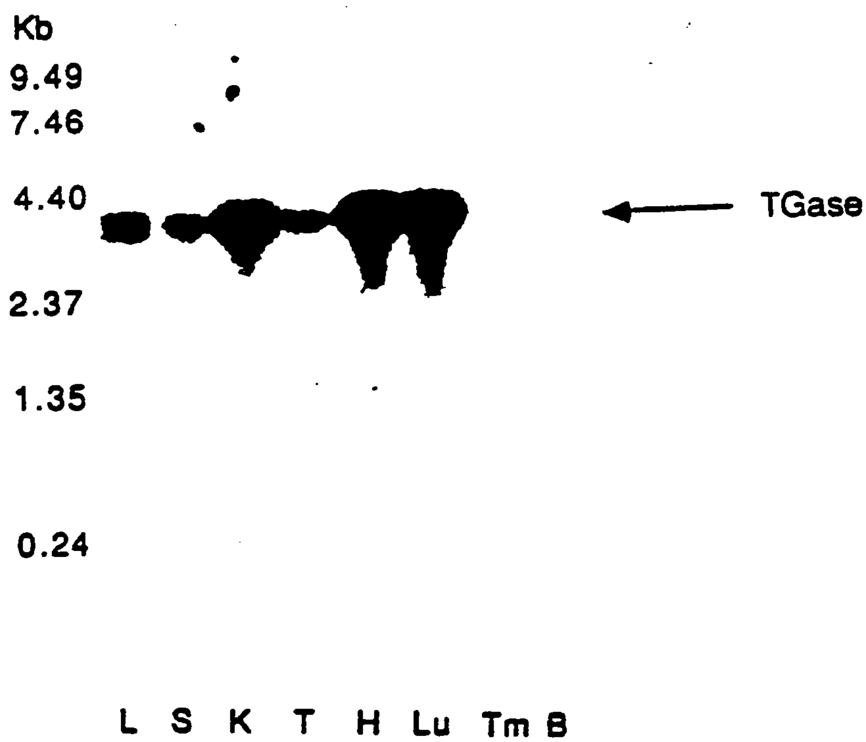


Fig. 5

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US91/09784

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(5): C12N 9/10, 15/54, 15/62, 15/80, 15/85
U.S.Cl.: 435/69.1, 193, 252.3, 320.1; 536/27; 935/8, 14, 68, 70

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
U.S.Cl.	435/69.1, 193, 252.3, 320.1; 536/27; 935/8, 14, 68, 70

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

APS, STN/Chemical Abstracts and DIALOG/Biosis databases

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
X,O Y	THE JOURNAL OF CELL BIOLOGY, Vol. 109, No. 4 Part 2, issued 9 November 1989, Gentile et al., "Isolation and Characterization of cDNA and Genomic Clones of Human Endothelial Cell Transglutaminases", page 198a, see abstract 1068.	1-5, 19-20 10-17
X,O Y	FASEB JOURNAL, Vol. 3, No. 4, issued 17 February 1989 Saydak et al., "cDNA Cloning of Mouse Macrophage Tissue Transglutaminase", page A1209, see abstract 5708.	6-9, 19-20 10-17
X,O Y	2nd INTERNATIONAL CONFERENCE ON TRANSGLUTAMINASES & PROTEIN CROSS-LINKING REACTIONS, 24-28 June 1990, Gentile et al., "Molecular Cloning and Sequence Analysis of cDNA for a Retinoic Acid-Inducible Tissue Transglutaminase from Human Umbilical Vein Endothelial Cells", see abstract.	1-5, 10, 14-16, 19-20 6-9, 11-13, 15-18
Y	US, A, 4,929,554, GOEDDEL et al., 29 May 1990, see columns 17-21, figures 7-8 and 11-16.	10-17
Y	EP, A, 0,268,772, DAVIE et al., 01 June 1988, see pages 16-20, Figures 5-11.	10-18

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
26 March 1992	20 APR 1992
International Searching Authority	**Signature of Authorized Officer**
ISA/US	William W. Moore

Form PCT/ISA/210 (second sheet) (Rev.11-87)

BNSDOCID: <WO 9212238A1 I >